Grenze International Journal of Engineering and Technology, June Issue

A Survey of Thumbnail Image Reconstruction Techniques: A Forensic Perspective

Lavanya M C[1] and Trisiladevi C. Nagavi[2]
[1-2]Department of CSE, JSSSTU, Mysuru
Email: lavanya.mc23@gmail.com, trisiladevi@sjce.ac.in

***Abstract*— Thumbnail images are small previews generated by operating systems, mobile gallery applications, image viewers, cameras, and image-processing software. In digital forensics, thumbnails are important as they remain available even when the original image is deleted, moved, overwritten, or stored on external media no longer attached to the device. In image reconstruction research, thumbnails represent a difficult low-resolution input: they are usually down-sampled, cropped, compressed, and stripped of much of the high-frequency content required for faithful high-resolution recovery. This work reviews the literature on thumbnail image reconstruction techniques ranging from direct recovery and classical interpolation to CNN-based super-resolution, GAN/perceptual methods, super-resolution, transformer-based restoration, diffusion-based super-resolution. The review presented in this paper includes identifying a research gap with respect to metadata-aware, uncertainty-aware, and hallucination-aware thumbnail reconstruction systems designed to focus various forensic review.**

***Keywords*— Thumbnail forensics; image reconstruction; super-resolution; SRCNN; ESRGAN; diffusion models; uncertainty estimation; hallucination-aware reconstruction; digital evidence.**

I. INTRODUCTION

Digital images are among the most common artefacts examined in modern digital investigations. However, the original image file may not always be available. One might have deleted it, overwritten, transferred to removable media, removed by a cloud synchronization or specific application workflow. Thumbnail artefacts are therefore important because they can preserve a reduced visual representation.

The term thumbnail reconstruction refers to recovering thumbnail artefacts from sources such as caches, metadata, application folders, or deleted files, and linking them to their possible origins. Improving a low-resolution thumbnail into a clearer or higher-resolution image using interpolation, restoration models, or super-resolution techniques. A rigorous study of thumbnail reconstruction should therefore consider both recovery and verification. This distinction is important because a reconstructed image may appear visually convincing even when the model infers details that are not directly supported by the original thumbnail.

Thumbnails are small previews created to speed-up browsing, indexing, previewing, or searching. It is generated by an operating system when a folder is opened, by a mobile gallery application when images are stored, by a camera as an embedded JPEG, or by third-party image viewers. The thumbnail-generation process involves various stages such as cropping, filtering, down-sampling, contrast or brightness adjustment, and JPEG compression [6], [7]. These transformations remove information and introduce compression artefacts.

Grenze ID: 01.GIJET.12.2.14_25

However, they also preserve enough visual evidence to identify broad scene content, faces, documents, contraband categories, logos, or the approximate identity of an image. Different platforms create thumbnails under different triggers, so user-action interpretation must be platform-specific [1]-[5].

II. THUMBNAIL RECONSTRUCTION TECHNIQUES

In reconstruction process the low-resolution thumbnail is processed to obtain a clearer or higher-resolution representation. However, in forensic applications, the reconstructed output should not be treated as the original image. These reconstruction techniques must be evaluated not only by visual quality, but also by explainability, reproducibility, and risk of hallucinating unsupported details.

A. Direct recovery and evidential reconstruction

One of the first forensically conservative technique is the direct recovery. It demonstrates the extraction of thumbnail exactly as stored, decode it with no enhancement, hash it, preserve the source container, and report its provenance. This method does not increase visual resolution, but it is the most defensible evidence because the output is directly observed. In forensic reporting, the original extracted thumbnail should always be retained alongside any enhanced or reconstructed output.

B. Metadata-aware inverse thumbnail modelling

A more advanced approach is to model how the thumbnail was generated. Farid presents the thumbnail-generation pipeline as a sequence of crop, filtering, down-sampling, post-filtering, brightness/contrast adjustment, and JPEG compression operations [6].

C. Classical interpolation

Classical interpolation methods such as bilinear, bicubic, nearest-neighbour and Lanczos upscaling. These methods are simple, deterministic, fast, and explainable. They do not hallucinate new objects, but they also cannot recover missing high-frequency details [9].

D. CNN-based single-image super-resolution

SRCNN introduced a direct end-to-end mapping from a low-resolution image to a high-resolution image and showed that sparse-coding-inspired super-resolution could be represented through a CNN-like pipeline [8]. FSRCNN improved computational efficiency by operating directly in low-resolution space and introducing a compact architecture [9].
ESPCN is a efficient sub-pixel convolution method which demonstrates real-time single-image and video super-resolution by extracting features in low-resolution space before learned upscaling [10].

E. Perceptual and GAN-based super-resolution

SRGAN concentrates on perceptual quality rather than pixel-wise distortion by combining adversarial loss with perceptual/content losses, producing photo-realistic textures at large upscaling factors [11]. ESRGAN uses residual-in-residual dense blocks, relativistic adversarial learning, and improved perceptual loss, while noting that hallucinated details and artefacts are a challenge in GAN-based super-resolution [12]. Real-ESRGAN extended this direction to real-world blind super-resolution by modelling complex degradations synthetically, including ringing and overshoot artefacts [13].

F. Transformer and diffusion-based reconstruction

SwinIR introduced transformer-based restoration in super-resolution by using residual Swin Transformer blocks. It demonstrated strong results in comparison with classical, lightweight, real-world super-resolution [14]. Diffusion-based approaches such as SR3 is an iterative denoising/refinement process that can generate highly realistic high-resolution images [15]. StableSR later used pretrained text-to-image diffusion priors for real-world super-resolution [16].
For thumbnails, where the input may be as small as 96, 160, or 256 pixels in one dimension, the number of possible high-resolution originals is enormous. Therefore, a forensic diffusion-based reconstruction should be presented as one possible reconstruction, not as the recovered original image.

G. Uncertainty-Aware and Hallucination-Aware Reconstruction

A major limitation of standard super-resolution is that most outputs do not indicate where the model is uncertain. Ning et al. proposed uncertainty-driven loss for single-image super-resolution, arguing that texture and edge regions carry high visual importance and should be handled adaptively through uncertainty modelling [17]. Liu, Cheng, and Tan proposed spectral Bayesian uncertainty for image super-resolution, estimating uncertainty in both spatial and frequency domain, which is directly relevant to the reliability of high-frequency reconstructed details [18]. More recent work has applied Monte Carlo dropout and deep ensembles to ESRGAN-style models to provide predictive uncertainty estimates [19].

Hallucination-aware reconstruction is even more directly relevant to forensic thumbnails. Recent work on hallucination scoring in generative image super-resolution argues that traditional metrics such as PSNR, SSIM, and LPIPS may not fully capture hallucinated visual elements that fail to match the low-resolution input or ground truth [20]. This is important because forensic users care not only about perceptual quality but about evidential fidelity. A forensic reconstruction system should therefore include hallucination warnings, uncertainty maps, and a clear visual separation between observed thumbnail pixels, deterministic enhancement, and model-inferred high-frequency detail.

Table 1 summarizes the literature with respect to thumbnail reconstruction and their key contributions:

TABLE I. PAPER REFERENCES AND THEIR KEY CONTRIBUTIONS

Reference No.	Paper Group	Key contributions
[1]	Morris and Chivers, Windows 7 thumbnail cache	Windows cache structure and behaviour
[2]	McKeown, Russell, and Leimich, centralized Windows thumbnail caches	Fast forensic triage using caches
[3]	Di Leom et al., Android thumbnails	Android thumbnail recovery and case study
[5]	van der Meer et al., image-viewer thumbnails	Automated extraction from image-viewer thumbnail databases
[8]-[10]	SRCNN, FSRCNN, ESPCN, VDSR, EDSR, RDN, RCAN	CNN/residual super-resolution foundations
[11]-[13]	SRGAN, ESRGAN, Real-ESRGAN	Perceptual/GAN and real-world blind SR
[14]-[16]	SwinIR, SR3, StableSR	Transformer and diffusion-based restoration
[17]-[21]	Uncertainty and hallucination-aware SR papers	Trustworthy enhancement for high-risk use

III. EVALUATION METRICS FOR THUMBNAIL RECONSTRUCTION

Thumbnail reconstruction can be evaluated using image quality and forensic validity. Peak Signal-to-Noise Ratio (PSNR) is a popular metric which measures pixel-level distortion. Structural Similarity Index Measure (SSIM) compares structural similarity and reflects visual structure better [19]. Learned Perceptual Image Patch Similarity (LPIPS) uses deep feature representations and generally correlates better with human perceptual similarity [20]. The perception-distortion trade-off formalizes the tension between low distortion and perceptual realism: methods that produce more realistic images may move farther from the exact ground truth [21].

TABLE II. OVERVIEW OF EVALUATION METRICS

Metric family	What it measures	Forensic interpretation
PSNR	Pixel-level distortion	Useful baseline for controlled experiments; may reward blurry outputs
SSIM / MS-SSIM	Structural similarity	Useful for comparing shape and luminance structure
LPIPS	Learned perceptual similarity	Useful for perceptual similarity, but not a forensic truth guarantee
FID / NIQE / MUSIQ	Distributional or no-reference quality	Useful for visual quality; weak for evidence fidelity
Uncertainty map	Where reconstruction is unstable	Essential for identifying unsupported fine detail
Hallucination score	Mismatch between generated details and input evidence	Important for GAN/diffusion outputs
Task-level checks	OCR, face matching, object detection	Should be treated as secondary, not proof of recovered ground truth

IV. A COMPARISON OF THUMBNAIL RECONSTRUCTION TECHNIQUES

Thumbnail reconstruction techniques vary in their assumptions, output quality, computational cost, interpretability, and forensic reliability. Classical methods are more easy to interpret and reproduce, whereas modern learning-based methods generate visually better reconstructions but can introduce hallucinations. Table 4 presents the summary of techniques that are useful in thumbnail construction from a forensic viewpoint.

TABLE III. SUMMARY OF THUMBNAIL RECONSTRUCTION TECHNIQUES

Technique family	Representative approach	Forensic defensibility
Direct extraction	Parse/cache carve thumbnail exactly as stored	Highest
Classical interpolation	Nearest, bilinear, bicubic, Lanczos	High
Metadata-aware inverse modelling	Estimate crop, scale, filtering, compression pipeline	High to medium

CNN/residual SR	SRCNN, VDSR, EDSR, RDN, RCAN	Medium
GAN/perceptual SR	SRGAN, ESRGAN, Real-ESRGAN	Low to medium
Transformer SR	SwinIR and related models	Medium
Diffusion SR	SR3, StableSR and related models	Low unless controlled
Uncertainty/hallucination-aware SR	Bayesian, ensemble, spectral uncertainty, hallucination scoring	Potentially high

V. Research Opportunities

The reviewed literature shows a gap between forensic thumbnail artefact research and computer-vision super-resolution research. Forensic studies explain where thumbnails reside, how they persist, and how they can be used to infer provenance. Super-resolution studies explain how to produce visually sharper images. Reconstruction quality is often discussed operationally, but less often with formal uncertainty reporting.

The strongest forensic use of thumbnail reconstruction is in investigations, visual inspections, and comparison against known images. AI enhancement techniques for thumbnails lack standardized forensic validation protocols and courtroom-ready reporting frameworks.

VI. Conclusion

Thumbnails can persist even after deletion, device reset under operating system previews, and embedded metadata. The review presented in this paper covers a comprehensive survey on the super-resolution techniques. The paper summarizes and compares various methods for visual enhancement, starting from interpolation and CNNs to GANs, transformers, and diffusion models. In addition, this work also presents a forensic point of view which requires more than visual quality. It requires provenance, reproducibility, uncertainty quantification, and hallucination control. The paper also proposes a promising research direction that can include metadata-aware, uncertainty-aware, and hallucination-aware thumbnail reconstruction layers. Such a framework can improve visual interpretability while preserving a clear boundary between observed evidence and model-generated inference.

References

[1] S. Morris and H. Chivers, "An analysis of the structure and behaviour of the Windows 7 operating system thumbnail cache," Proceedings of the 1st Cyberforensics Conference, 2011.

[2] S. McKeown, G. Russell, and P. Leimich, "Fast Forensic Triage Using Centralised Thumbnail Caches on Windows Operating Systems," Journal of Digital Forensics, Security and Law, vol. 14, no. 3, 2019.

[3] M. D. Leom, C. J. D Orazio, G. Deegan, and K.-K. R. Choo, "Forensic Collection and Analysis of Thumbnails in Android," IEEE Trustcom/BigDataSE/ISPA, 2015, pp. 1059-1066.

[4] B. Findlay, "A review of thumbnail images artefacts in the Linux desktop and a methodology to add provenance to deleted files," Forensic Science International: Digital Investigation, vol. 44, 2023.

[5] Abdullah, R. Ibrahim, and K. Mohamad, "Carving Thumbnail/s and Embedded JPEG Files Using Image Pattern Matching," Journal of Software Engineering and Applications, vol. 6, no. 3B, pp. 62-66, 2013.

[6] C. Dong, C. C. Loy, K. He, and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," IEEE TPAMI, vol. 38, no. 2, pp. 295-307, 2016.

[7] C. Dong, C. C. Loy, and X. Tang, "Accelerating the Super-Resolution Convolutional Neural Network," ECCV, 2016.

[8] W. Shi et al., "Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network," CVPR, 2016.

[9] C. Ledig et al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," CVPR, 2017.

[10] X. Wang et al., "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks," ECCV Workshops, 2018.

[11] X. Wang, L. Xie, C. Dong, and Y. Shan, "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data," ICCV Workshops, 2021.

[12] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, "SwinIR: Image Restoration Using Swin Transformer," ICCV Workshops, 2021.

[13] C. Saharia et al., "Image Super-Resolution via Iterative Refinement," IEEE TPAMI / arXiv, 2021.

[14] J. Wang, Z. Yue, S. Zhou, K. C. K. Chan, and C. C. Loy, "Exploiting Diffusion Prior for Real-World Image Super-Resolution," IJCV, 2024.

[15] Q. Ning et al., "Uncertainty-Driven Loss for Single Image Super-Resolution," NeurIPS, 2021.

[16] T. Liu, J. Cheng, and S. Tan, "Spectral Bayesian Uncertainty for Image Super-Resolution," CVPR, 2023.

[17] M. S. Adapa et al., "Uncertainty Estimation for Super-Resolution using ESRGAN," arXiv, 2024.

[18] W. Ren et al., "Hallucination Score: Towards Mitigating Hallucinations in Generative Image Super-Resolution," arXiv, 2025.

[19] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, 2004.
[20] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric," CVPR, 2018.
[21] Y. Blau and T. Michaeli, "The Perception-Distortion Tradeoff," CVPR, 2018.